DOCUMENT RESUME

ED 466 650 TM 034 261

AUTHOR	Bishop, John H.
TITLE	What Should Be the Federal Role in Supporting and Shaping Development of State Accountability Systems for Secondary School Achievement?
SPONS AGENCY	Office of Vocational and Adult Education (ED), Washington, DC.
PUB DATE	2002-04-00
NOTE	44p.
CONTRACT	ED-99-CO-0160
PUB TYPE	Opinion Papers (120) -- Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Accountability; Educational Change; *Federal Government; *Government Role; National Competency Tests; Secondary Education; *State Programs; Test Construction; Test Use; Testing Programs
IDENTIFIERS	*Exit Examinations; *External Examination Program

ABSTRACT

This paper explores the federal role in the development of state accountability systems for secondary school achievement. The first section of the paper examines seven suggested proximate causes for the poor performance of U.S. secondary school students. This section concludes that the causes for the relatively poor performance of U.S. secondary school students are the poor quality of teachers, the academic standards set by teachers and administrators, and the culture of secondary schools. The second section of the paper proposes an institutional mechanism for raising standards and improving student engagement and motivation: curriculum-based external exit examinations (CBEEES). Studies of the impact of CBEEES have found that they improve teaching and increase learning. Section 3 describes the strategies that state governments in the U.S. have devised to reform secondary education. Section 4 summarizes research on the effects of these strategies. This research suggests that CBEEES are the most effective of the strategies being tried, although stakes for schools are also effective. High school graduation tests do not appear to have big effects on test scores when other standards-based reforms are controlled. They do, however, have big effects on employer perceptions of the competence of recent high school graduates and on the wages and earnings of these graduates. (Contains 3 tables, 24 endnotes, and 17 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 466 650

What Should Be The Federal Role in Supporting and Shaping Development of State Accountability Systems for Secondary School Achievement?

John H. Bishop

Cornell University

Department of Human Resource Studies

April 2002

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This paper was prepared for the Office of Vocational and Adult Education, U.S. Department of Education pursuant to contract no. ED-99-CO-0160. The findings and opinions expressed in this paper do not necessarily reflect the position or policies of the U.S. Department of Education.

TM034261

BEST COPY AVAILABLE

What Should Be The Federal Role in Supporting and Shaping Development of State Accountability Systems for Secondary School Achievement?

John H. Bishop

Introduction

There is much to be proud of in American education. Nearly 30 percent of the nation's youth now obtain a four-year college degree. The graduates of American universities have generated many of the major technological breakthroughs of the last quarter century. Primary education is also quite successful. In recent international assessments fourth graders in the U.S. placed number two in reading literacy, number three in science and number twelve (out of 26) in mathematics.

Secondary education, however, is a different story. In the 1960s U.S. participation rates in secondary education were the highest in the world. This is no longer true. According to the OECD data presented in Table 1, enrollment rates of 16 and 17 year olds in Australia, Belgium, Canada, Denmark, Finland, France, Germany, Japan, Korea, the Netherlands, Norway and Sweden all exceed U.S enrollment rates by 10 percentage points or more.[1] Graduation rates are also higher in these countries.

The rate at which U.S. students learn new skills clearly decelerates during secondary school. Gains on the TIMSS math and science assessments from 4th to 8th grade are smaller for the US than any other country [see columns 5 and 6 of Table 1]. The IEA Study of Reading Literacy had similar findings [see column 7].[2] In the reading literacy study American students fell from their number two spot in fourth grade to 14th amongst 24 rich industrialized countries in ninth grade.[3] The most telling indicator of the poor quality of American secondary schools is the TIMSS results for students at the end of secondary school (see column 9 and 10 of Table 1). In mathematics seniors in U.S. high schools ranked 19th out of 21 nations, ahead of only Cyprus and South Africa. In science U.S. seniors ranked 16th out of 21, ahead of Cyprus, Italy, Hungary, Lithuania and South Africa.

How do students who lead the world in 4th grade get transformed into cellar dwellers at the end of upper secondary school? In the first section of the paper I examine seven proposed proximate causes of the poor performance of U.S. secondary schools. I conclude that spending less money or spending less time in school is not responsible for our lag behind European competitors. Rather the causes appear to be the quality of teachers, the academic standards set by teachers and administrators and the culture of secondary schools. The second section of the paper proposes an institutional mechanism for raising standards and improving student engagement and motivation: curriculum-based external exit examinations (CBEEES). Studies of the impacts of CBEEES have found that they improve teaching and increase learning. Section 3 describes the strategies that state governments in the U.S. have devised to reform secondary education. Section 4 presents a summary of research my colleagues and I have conducted evaluating the effects of these strategies. We have concluded that curriculum-based external exit exams are the most effective of the strategies being tried. Stakes for schools--rewarding schools that improve student performance and sanctioning schools that fail to meet targets for student achievement--are also effective. High school graduation tests (minimum competency exams that must be passed to receive a high school diploma) do not appear to have big effects on test scores when other standards-based reforms are controlled. They do, however, have big effects on employer perceptions of the competence of recent high school graduates and on the wages and earnings of these graduates.

The final section of the paper discusses the policy choices facing states and the U.S. Department of Education. It provides guidance for writing regulations for the "No Child Left Behind" Act and proposes a modest federal investment in merit scholarships and other programs designed to improve school culture, teaching standards and student incentives to learn.

The Proximate Causes of the Poor Performance of American Secondary Schools: Teacher Quality, Student Engagement and School Culture

We begin by examining the proximate causes of low achievement at the end of secondary school. The discussion is organized around seven topics--each of them a proposed explanation for the poor performance of U.S. students relative to their counterparts in northern Europe and East Asia.

1) Teacher quality and compensation
2) Expenditure per pupil
3) Time devoted to instruction and study
4) Engagement--Effort per unit of scheduled time
5) Nerd Harassment—Peer Pressure against Studiousness
6) Students Avoiding Rigorous Courses
7) Pressures on Teachers to Lower Standards

Teacher Quality and Compensation

Teacher quality has big effects on student learning. The teacher's general academic ability and subject knowledge are the characteristics that most consistently predict student learning (Hanushek 1971, Strauss and Sawyer 1986, Ferguson 1990, Ehrenberg and Brewer 1993, Monk 1992).

Unfortunately, teaching secondary school does not attract the kind of talent that is attracted into the profession in Europe and East Asia. In 1999-2000 intended education majors had SAT scores that were 33 points below average in mathematics and 22 points below average on the verbal test (NCES 2000, Table 135). School administrators are also remarkably willing to hire and assign staff to teach subjects that are outside their field of expertise and training. Teachers who neither majored nor minored in history in college teach more than half of secondary school history classes. Teachers who did not major or minor in a physical science or engineering in college teach more than half of chemistry and physics students.[4]

Recent college graduates recruited into math or science teaching jobs spent only 30 percent of their college career taking science and mathematics courses. Since 46 percent had not taken a single calculus course, the prerequisite for most advanced mathematics courses, it appears that most of the math taken in college was reviewing high school mathematics (NCES 1993b, p. 428-429). The graduates of the best American universities typically do not enter secondary school teaching because the pay and conditions of work are relatively poor.

Despite the fact that wage rates and standards of living in the U.S. are higher than in any other OECD nation, there are six countries—Australia, Germany, Japan, Korea, Switzerland and the United Kingdom—that have higher annual salaries for secondary school teachers (see column 11 of Table 1). Comparisons of secondary school teacher salaries with per capita GDP are presented in column 12. American upper secondary teachers with 15 years of experience are paid only 10 percent more than the nation's per capita GDP. In Europe and East Asia by contrast salaries for teachers with 15 years of experience are on average 65 percent higher than per capita GDP (OECD, 2000, p. 215).

The lower pay in the United States is not a tradeoff for more attractive conditions of work. Indeed the working conditions of U.S. secondary school teachers are considerably less attractive. Their contracted teaching hours are 954 hours per year on average; 50 percent more then the mean for the other OECD nations in the table--635 hours (OECD, 2000, p. 229). When you divide their annual salaries by the contracted number of teaching hours, lower secondary school teachers with 15 years of experience are paid only $34.00 per hour. The average for the other OECD countries is $47.66, forty percent more (OECD, 2000, p. 16). In other occupations hourly wages are higher in the US. Why

do we pay our secondary school teachers so little? Is standards based reform likely to improve the qualifications and pay of teachers? These questions are taken up later in the paper.

School Expenditures

When expenditures per secondary school student are deflated by a purchasing power parity price index, the U.S. spends more than other countries with sole exception of Switzerland. However, teachers of constant quality are more expensive in America than in Europe and East Asia because college graduates (the pool of workers from which teachers must be drawn) are better paid. Since labor compensation is the bulk of education costs, the proper deflator for schooling expenditure is not a general cost of living index, but a wage index that reflects among other things the cost of recruiting competent teachers. Lacking such an index, deflation by GDP per capita is the next best thing. OECD's latest estimates of the ratio of per pupil spending for secondary schools to per capita GDP are given in column 15 of Table 1. By this indicator most countries are pretty similar. The U.S. secondary school spending ratio is 7.4 percent below the average for the other nations in the table (OECD, 2000, p. 95).

How is it possible for the U.S. to pay its teachers so little and yet end up spending so much on secondary education? Japan and Korea keep per pupil costs down by increasing class size substantially above U.S. levels. Europe, however, does not. Pupil teacher ratios in Europe and the U.S. are very similar. What's happening to the money saved by paying American teachers low hourly wages? It's being used to provide a variety of non-instructional services such as after-school sports, bus transportation, psychological counseling, medical check ups, after-school day care, hot meals, and driver education that other countries typically assign to other institutions. In Japan and Europe students use public transportation to commute to school, so transportation is not charged to the school budget. In many European countries, local governments, not schools, sponsor after-school sports programs. These additional functions of American schools require extra non-teaching staff. Non teachers account for 22 percent of current expenditure on K-12 education in the US; only 14 percent of current expenditure in other OECD nations (see column 16 of Table 1).[5] If adjustments were made for service mix and a cost-of-education index reflecting compensation levels in alternative college-level occupations were used to deflate expenditure, the U.S. advantage in instructional spending per pupil would drop.

Time Devoted to Instruction

Many studies have found learning to be strongly related to time on task (Wiley 1986, Walberg 1992). OECD estimates of annual hours of instruction for 14-year-old students are presented in column 9 of Table 1. These numbers contradict the widely held belief that U.S. students do poorly because of shorter school days and shorter school

years. Only 5 of the OECD countries in the table assign their students to attend classes for more hours per year than the United States. Twelve countries have their 14 year olds in school for less time. **Why does an hour of instruction in European and East Asian classrooms produce more learning than in American classrooms?**

Engagement--Effort per Unit of Scheduled Time

Classroom observation studies reveal that American students actively engage in learning activities for only about half the time they are scheduled to be in a classroom. A study of schools in Chicago found that public schools with high-achieving students averaged about 75 percent of class time for actual instruction; for schools with low achieving students, the average was 51 percent of class time (Frederick, 1977). Overall, Frederick, Walberg and Rasher (1979) estimated 46.5 percent of the potential learning time is lost due to absence, lateness, and inattention.

Just as important as the amount of time participating in a learning activity is the intensity of the student's involvement in the process. The high school teachers surveyed by John Goodlad (1983) ranked "lack of student interest" as the most important problem in education and "lack of parent interest" as the second most important problem. Why is student engagement so low? Poor teaching possibly, but there are other explanations as well.

Nerd Harassment

Probably the most important reason for lack of student engagement in the U.S. is a peer culture that is often hostile to studiousness and public displays of enthusiasm for academic learning. Twenty four percent of the 95,000 secondary school students recently surveyed by the Educational Excellence Alliance said "My friends make fun of people who try to do well in school." Interviews I conducted of middle school boys in Ithaca New York in 1996 and 1997 revealed that most of them internalized a norm against "sucking up" to the teacher. How does a boy avoid being thought a "Suck up?" He:

- Avoids giving the teacher eye contact
- Does not hand in homework early for extra credit,
- Does not raise his hand in class too frequently, and
- Talks or passes notes to friends during class (signaling that you value friends more than your rep with the teacher).

Similarly, Steinberg, Brown and Dornbusch's recent study of nine high schools in California and Wisconsin concluded that:

> ... less than 5 percent of all students are members of a high-achieving crowd that defines itself mainly on the basis of academic excellence... Of all the crowds the 'brains' were the least happy with who they are--nearly half wished they were in a different crowd.[6]

Why are the studious called ***suck ups***, ***dorks*** and ***nerds*** or accused of "***acting white***"? Why are students who disrupt the class or try to get the class off track, not sanctioned by their classmates? In part, it is because many teachers grade on a curve and this means trying hard to do well in a class is making it more difficult for others to get top grades. When exams are graded on a curve or college admissions are based on rank in class, joint welfare is maximized if no one puts in extra effort. In the repeated game that results, side payments--friendship and respect--and punishments—ridicule, harassment and ostracism--enforce the cooperative "don't study much, hang out instead" solution. If, by contrast, students were evaluated relative to an outside standard, they would no longer have a personal interest in getting teachers off track or persuading each other to refrain from studying. Peer pressure demeaning studiousness might diminish. We will return to this issue later in the paper.

Student Preference for Easy Courses

Although research has shown that learning gains are substantially larger when students take honors and AP courses,[7] enrollment in these courses is quite limited. In many schools guidance counselors allow only a select few into these courses. Many students prefer easy courses. In the 1987 survey, 62 percent of 10th graders agreed with the statement, ***"I don't like to do any more school work than I have to."***[8] Parents often agree with their child. As one guidance counselor described:

> ***A lot of... parents were in a 'feel good' mode."...If they [the students] felt it was too tough, they would back off. I had to hold people in classes, hold the parents back. [I would say] "Let the kid get C's. It's OK. Then they'll get C+'s and then B's." [But they would demand,] "No! I want my kid out of that class!"*** [9]

Rigorous courses are avoided because the rewards for the extra work are small for most students. While selective colleges evaluate grades in the light of course demands, many colleges have, historically, not factored the rigor of high school courses into their admissions decisions. Trying to counteract this problem, college admissions officers have been telling students that they are expected to take the most rigorous courses offered by their school. This effort has met with some success. More students are taking chemistry and physics and advanced mathematics. But many students have not gotten the message and still think taking easy courses is a good strategy. One student told a reporter:

> *My counselor wanted me to take Regents history and I did for a while. But it was pretty hard and the teacher moved fast. I switched to the other history and I'm getting better grades. So my average will be better for college.*[10]

Consequently, the bulk of students who do not aspire to attend selective colleges quite rationally avoid rigorous courses and demanding teachers.

Pressure on Teachers to Lower Standards

When teachers try to set high standards, they often get pressured to go easy. Thirty percent of American teachers say they "feel pressure to give higher grades than students' work deserves." Thirty percent also feel pressured "to reduce the difficulty and amount of work you assign."[11] Students also pressure teachers to go easy. Sizer's description of Ms. Shiffe's biology class, illustrates what sometimes happens:

> She wanted the students to know these names. They did not want to know them and were not going to learn them. Apparently no outside threat--flunking, for example--affected the students. Shiffe did her thing, the students chattered on, even in the presence of a visitor....Their common front of uninterest probably made examinations moot. Shiffe could not flunk them all, and, if their performance was uniformly shoddy, she would have to pass them all. Her desperation was as obvious as the students' cruelty toward her. (1984 p. 157-158)

Some teachers are able, through the force of their personalities, to induce their students to undertake tough learning tasks. But for all too many, academic demands are compromised because the bulk of the class sees no need to accept them as reasonable and legitimate. Why are American students more interested in diplomas than in learning? Why are rewards for learning so weak? Why do school administrators assign staff to teach subjects they did not study in college?

Weak Organic Accountability Systems as Ultimate Cause: External Examinations as standard Setters and a Way to Boosting the Rewards for Learning

Most of the problems listed above are not present in Northern Europe and East Asia. Why are standards higher there? Why are school administrators more focused on students' academic achievement? If citizens of Japan, Korea, Britain, Denmark, France, Germany, the Netherlands and a host of other countries were asked these questions, they would point to their nation's system of curriculum-based external exit examinations (CBEEES). These examinations systems provide a strong and organic system of accountability. High stakes are attached to how students do on these exams. Exam

grades appear on resumes and are requested on job applications. Exam grades influence (and in some nations completely determine) whether a student can enter a university and which university and what field of study they are admitted to. In the United States, by contrast, admission to the best colleges depends on teacher assessments of relative performance--rank in class and grades--and multiple choice format aptitude tests that are not keyed to the courses taken in secondary school. Employers pay little attention to achievement in high school when making hiring decisions. Clearly CBEEES strengthen student incentives to study. Students are no longer competing with each other for a limited number or As and Bs. Everyone in the class can get a 90 or better on the external exam, so students will be less supportive of those who disrupt the class and more supportive of those who take learning seriously. It no longer makes sense for students to avoid the more rigorous courses and the more demanding teachers.

CBEEES fundamentally change how student achievement is signaled. By doing so they organically transform the incentives for everyone: parents, teachers and secondary school administrators as well as students. In the U.S. local school administrators serving at the pleasure of locally elected school boards make the thousands of decisions that determine academic expectations and program quality. When there is no external assessment of academic achievement, students, parents and local taxpayers benefit little from administrative decisions that opt for higher standards, more qualified teachers or a heavier student work load. The immediate consequences of such decisions are all negative: higher local property taxes, more homework, having to repeat courses, lower GPA's, complaining parents and a greater risk of being denied a diploma.

College admission decisions are based on rank in class, GPA and aptitude tests, not externally assessed achievement in secondary school courses, so upgraded standards will not improve the college admission prospects of next year's graduates. Graduates will probably do better in difficult college courses and will be more likely to get a degree, but that benefit is uncertain, far in the future and not visible to voters in school board elections. In this environment, administrators will seek teachers who keep their class orderly and entertained, who have roots in the community and who are willing to coach. If this is all one expects of teachers, sufficient numbers can be found at current salary levels. If, however, administrators were to demand that newly hired teachers have a deep knowledge of their subject and the ability to teach it to teenagers, they would find that there are not enough qualified teachers to go around. The shortage would not disappear until much higher salaries were offered. External exams make stake holders care about how well high school subjects are taught. Hiring better teachers and improving the school's science laboratories now yields a visible payoff--more students passing the external exams and being admitted to top colleges. This should induce school districts to compete for talent by offering higher salaries and better working conditions.

When external assessment is absent, school reputations are determined largely by school characteristics over which teachers and administrators have no control: the socio-economic status of the student body and the proportion of graduates going to college. Consequently, higher standards do not benefit students as a group, so parents as a group have little incentive to lobby for higher teacher salaries, higher standards and higher school taxes. Under a system of external exams, teachers and local school administrators lose the option of lowering standards to reduce failure rates and raise self-esteem. The only response open to them is to demand more of their students so as to maximize their chances of being successful on the external exams.

External assessment of accomplishment puts students, teacher and parents on the same team. It assists the development of mentoring relationships between teachers and students. In the absence of external assessment, the effort to become friends with one's students and their parents tends to deteriorate into extravagant praise for mediocre accomplishment. In courts of law, judges must disqualify themselves when a friend comes before the bar. Yet, American teachers are placed in this double bind every day. Often the role conflict is resolved by lowering expectations. Other times the choice of high standards means that close supportive relationships are sacrificed.

A further benefit of CBEEES is the professional development that teachers receive when they come to centralized locations to grade the extended answer portions of examinations. In May 1996 I interviewed a number of teachers union activists about the examination system in the Canadian province of Alberta. Even though the union and these teachers opposed the exams, they universally reported that serving on grading committees was "...a wonderful professional development activity (Bob, 1996)." Having to agree on what constituted excellent, good, poor, and failing responses to essay questions or open ended math problems resulted in a sharing of perspectives and teaching tips that most found very helpful.

CBEEES should, consequently, influence the resources made available to schools, the priorities of school administrators, teacher pedagogy, parental for schools and student effort.

Careful empirical analysis of data from the Third International Mathematics and Science Study (TIMSS and TIMSS-R) and the International Assessment of Educational Progress has found that teaching is more rigorous and students learn more in nations with CBEEES.[12] Thirteen-year-old students from countries with CBEEE systems outperform students from other countries at a comparable level of economic development by **.67 to 2.0** grade level equivalents (GLE) in mathematics, science, geography and reading literacy. Closer to home, students in Canadian provinces with diploma exams were a statistically significant .5 GLE ahead in math and science of comparable students in other provinces

The impacts of CBEEES on school policies and instructional practices have also been studied. CBEEES are associated with higher minimum standards for becoming a teacher, higher teacher salaries (30-34 percent higher for secondary school teachers) and a greater likelihood of hiring teachers who have majored in the subject they are assigned to teach and specialize in teaching it. Schools in CBEEES jurisdictions equip better science labs, devote more hours to math and science instruction and provide after school tutoring to more students.

Fears that CBEEES have caused the quality of instruction to deteriorate appear to be unfounded. Students in CBEEES jurisdictions were less likely to say that memorization is the way to learn the subject and more likely to do experiments in science class. Quizzes and tests were more common, but in other respects pedagogy was no different. They were no less likely to like the subject and they were more likely to agree that "science is useful in every day life." Students also talked with their parents more about schoolwork and reported their parents had more positive attitudes about the subject.

What do these positive findings regarding the organic accountability effects of curriculum-based external exit exams in other countries suggest about how our standards based reform efforts should be structured?

STANDARDS-BASED REFORM

American policy makers are trying to deal with the low standards and weak incentives for hard study by making students, staff and schools more accountable for learning. The education departments of the 50 states have responded by developing content standards for core academic subjects, administering tests assessing this content to **all** students, publishing individual school results and holding students and schools accountable for student achievement. While these efforts are generically referred to as standards-based reform, the mix of initiatives varies a great deal from state to state.

Domestic Curriculum-Based External Examination Systems

While many states--Maryland, Georgia, Mississippi, Oklahoma, Arkansas, Tennessee, Texas, Virginia, Michigan, etc.—are developing end-of-course exams for key high school subjects and appear to be planning to implement a CBEEES, only two states—New York and North Carolina—actually had one during the 1990s. State sponsored systems of end-of-course exams are described in Table 2. The grand daddy of these examination systems is New York's Regents exam system. It has been in continuous operation since the 1860s. Panels of local teachers grade the exams using rubrics supplied by the state Board of Regents. Exam scores appear on transcripts and

are the final exam mark that is averaged with the teacher's quarterly grades to calculate the final course grade. A college bound student taking a full schedule of Regents courses would typically take Regents exams in mathematics and earth science at the end of 9th grade; mathematics, biology and global studies exams at the end of 10th grade; mathematics, chemistry, American history, English and foreign language exams at the end of 11th grade and a physics exam at the end of 12th grade. However, taking Regents courses and therefore Regents exams was voluntary until late in the 1990s. Prior to 1998 nearly half of students chose to take 'local' courses intended originally for non-college bound students and where good grades could be obtained without much effort.

North Carolina introduced end-of-course exams for Algebra 1 and 2, Geometry, Biology, Chemistry, Physics, Physical Science, American History, Social Science and English 1 between 1988 and 1991. Other versions of these courses not assessed by a state test do not exist, so virtually all North Carolina high school students take at least six of these exams. Test scores appear on the student's transcript and most teachers have been incorporating EOC exam scores in course grades. Starting in the year 2000, state law requires the EOCE tests to have at least a 25% weight in the final course grade. Clearly from this description one can see that North Carolina's end-of-course exams and New York's Regents Exams prior to 1999 carried low to moderate stakes for students, not high stakes.

Most states pursuing standards based reform have established test based school accountability systems and high stakes minimum competency high school graduation exams (MCEs) that are quite different from CBEEES.

Minimum Competency Graduation Exams

Eighteen states have minimum competency exam graduation requirements applying to the graduating class of 2000. Another eleven states are developing or phasing in MCEs. MCEs raise standards, but probably not for everyone.[13] The standards set by the teachers of honors classes and advanced college prep classes are not changed by an MCE. Students in these classes pass the MCE on the first try without special preparation. The students who are in the school's least challenging courses experience the higher standards. Students pursuing the "Do the Minimum" strategy are told "you must work harder" if you are to get the diploma and go to college. School administrators want to avoid high failure rates, so they are likely to focus additional energy and resources on raising standards in the early grades and improving the instruction received by struggling students.

School Report Cards and Stakes for Teachers and Administrators

So far we have discussed mechanisms for holding students accountable for learning. Formal systems for holding schools accountable are growing in popularity. In 1999 thirty-seven states were publishing school report cards for all or almost all of their schools.[14] Publicly identifying low performing schools is intended to spur local school administrators and boards of education to undertake remedial action. Nineteen states had a formal mechanism for rewarding schools either for year-to-year gains in achievement test scores or for exceeding student achievement targets.[15] Nineteen states had special assistance programs to help failing schools turn themselves around. If improvements were not forthcoming, eleven states had the power to close down, take over or reconstitute failing schools.

Exactly how are domestic student and school accountability strategies similar to or different from the CBEEES that are found abroad and in New York and North Carolina? We begin by noting the features they have in common. Minimum competency exams:

1. **Produce signals of accomplishment that have real consequences for students and schools.** While some stakes are essential, high stakes may not be necessary. Analyses of Canadian and US data summarized below suggest that moderate stakes may be sufficient to produce substantial increases in learning.

2. **Cover all or almost all students.**

3. **Define achievement relative to an external standard, not relative to other students in the classroom or the school.**

4. **Assess a major portion of what students are expected to know and be able to do.** Studying to prepare for an exam (whether set by one's own teacher or by a state department of education) should result in the student learning important material and developing valued skills. Some MCEs, CBEEES and teacher exams do a better job of achieving this goal than others. External exams, however, cannot assess every instructional objective. Teachers should be responsible for evaluating dimensions of performance that cannot be reliably assessed by external means or that local leaders want to add to the learning objectives specified by the state department of education.

5. **Are controlled by the education authority that establishes the curriculum for and funds K-12 education.** Curriculum reform is facilitated because coordinated changes in instruction and exams are feasible. Tests established and mandated by other organizations serve the interests of other masters. America's premier high stakes exams--the SAT-I and the ACT—serve the needs of colleges to sort students by aptitude, **not** the needs of schools to reward students who have learned what high schools are trying to teach.

Curriculum-based external exit exam systems are distinguished from MCEs by the following additional features. CBEEES:

1. **Signal multiple levels of achievement in the subject.** If only a pass-fail signal is generated by an exam and passing is necessary to graduate, the standard will almost inevitably to be set low enough to allow almost everyone to pass after multiple tries. This will not stimulate the great bulk of students to greater effort. CBEEES signal the student's achievement level in the subject, so all students, not just those at the bottom of the class, have an incentive to study hard to do well on the exam. Consequently, CBEEES should be more likely to improve classroom culture than a MCE.

2. **Assess more difficult material.** Since CBEEES are supposed to measure and signal the full range of achievement in the subject, they contain more difficult questions and problems. This induces teachers to spend more time on cognitively demanding skills and topics. MCEs, by contrast, are designed to identify which students have failed to surpass a rather low minimum standard, so they do not to ask questions or set problems that students near that borderline are unlikely to be able to answer or solve.[16] This tends to result in too much class time being devoted to practicing low-level skills.

3. **Are collections of End-of-Course Exams (EOCE).** Since they assess the content of specific courses, the teacher/s of that course (or course sequence) will inevitably feel responsible for how well their students do on the exam. Grades on EOCEs should be a part of the overall course grade further integrating the external exam into the classroom culture. Alignment between instruction and assessment is maximized and accountability is enhanced. Proponents argue that teachers will not only want to set higher standards, they will find their students more attentive in class and more likely to complete demanding homework assignments. They become coaches helping their team do battle with the state exam.

American Evidence on the effects of Standard Based Reform

Improvements in student performance on state exams are often cited as evidence that school accountability initiatives are working. Opponents disagree. Test scores have gone up, they say, because test preparation is displacing the teaching of other skills and knowledge that are more important to success in college and in jobs. This is a testable hypothesis. Bishop, Mane, Bishop and Moriarty (2001) and Bishop, Mane and Bishop (2000) have tested it by measuring the effects of accountability systems on college enrollment and labor market success after high school of a representative sample of eighth graders in 1988. We also measured impacts on academic achievement. To avoid teaching to the test effects we used achievement tests—the NAEP and NELS: 88 tests—

which are quite different from those used by the state accountability systems being evaluated.

States have introduced different packages of standards based reform initiatives, so we assessed their impacts by comparing outcomes in different states. We studied the impact of one old style reform—state mandated minimum course graduation requirements—and three different SBR policies:

1. Rewards for schools that improve on statewide tests and/or sanctions for failing schools—closure, reconstitution, loss of accreditation etc. [Since few states had implemented these policies prior to 1992, they are not included in our study of 1988 eighth graders]

2. Minimum competency exams

3. Curriculum-Based External Exit Exam System--i.e. the New York/North Carolina stakes for students policy mix during the 1990s.

The primary data set—NELS:88--provides six years of longitudinal data on 14,000 students who were 8th graders in 1988. Family background is a powerful predictor of high school completion, academic achievement, college attendance and labor market success, so our analyses included controls for a long list of socio-demographic characteristics of the student. We also controlled for the characteristics of the high school and the community—type of private school, teacher salary, pupil-teacher ratio, mean eighth grade test scores, ethnic and socio-economic composition of the student body, local unemployment rates, wage rates and the payoff to and tuition costs of college attendance. The eighth graders who subsequently dropped out of high school were tested and interviewed in 1992 and 1994 and so are included in the analysis sample.

Effects on College Attendance: Estimates of effects on the proportion of 8th graders who subsequently went to college are presented in Figure 1. The **s above a bar indicates that the outcome is significantly greater in MCE states at the 2.5 percent level. A * indicates significantly greater at the 5 percent level. A + above a bar indicates significantly greater at the 10 percent significance level. MCEs significantly increased the percentage of 8th graders who were attending college 6 years later (by 2.3 to 4.4 percentage points depending on GPA in 8th grade). CBEEES substantially increased college attendance rates of students with low GPAs in 8th grade. College attendance rates of high GPA students were unaffected.

Effects on Labor Market Success: Estimates of effects of exit exams on annual earnings are presented in Figure 2. Controlling on high school completion and college attendance, students who attended high school in states with MCEs earned significantly

more--9 percent more in the calendar year following graduation-- than students in states without MCEs.[17]

Effects on Test Scores: Our estimates of the effects of state imposed graduation requirements on scores on National Assessment of Educational Progress 8th grade assessments are summarized in Figure 3.[18] Estimates of the effect of graduation requirements on test score gains from 8th to 12th grade are presented in Figure 4.

The policy that clearly had the biggest effects on test scores was curriculum-based external exit examinations—the combination of EOCEs and MCEs that has been in place in New York State since the early 1980s and in North Carolina since about 1991. In comparison to students in states without MCEs or CBEEES, 8th graders in New York and North Carolina were about 45 percent of a grade level equivalent (GLE) ahead in math and science and 65 percent of a GLE ahead in reading. In addition, test score gains from 8th to 12th grade were nearly 40 percent of a grade level equivalent greater in New York State. This confirms and extends earlier findings that New York students did significantly better on SAT tests and the 1992 8th grade NAEP math tests than other states with demographically similar populations (Bishop, Moriarty and Mane 2000).

The next most powerful state policy was academic course graduation requirements. Students living in states that set academic course graduation requirements four units higher learn about one-third of a grade level equivalent more during high school.

The next most powerful SBR policy was stakes for teachers and schools particularly when rewards for successful schools were combined with sanctions for failing schools. The bars in Figure 3 depict our estimate of the effect of a state both rewarding schools for success and threatening to sanction failing schools. Students in these states were 20 percent of a GLE ahead in math and science of demographically comparable students in states that did neither. They were 24 percent of a GLE ahead in reading. Public reporting of school level results on state tests is necessary for the implementation of these policies, but on its own it had no discernable effect on student achievement.

When other SBR policies were held constant, the positive effects of state imposed MCEs on achievement were small and statistically insignificant. While state imposed MCEs had no significant effects on learning gains of students with average or above average grades in 8th grade, students with low GPAs learned more math and science when they lived in MCE states.

The policy having the smallest effects was state imposed elective and non-academic course graduation requirements. They had no effects on test score gains

during high school, no effects on earnings after high school and lowered college attendance rates.

Whose predictions were correct? Our analysis of college attendance rates, labor market success and test scores overwhelmingly rejects the hypotheses that test based accountability systems hurt students by inducing teachers to teach to severely flawed tests. Indeed the estimated impacts of test-based accountability policies on indicators of success after high school are positive, not negative as predicted by SBR critics. Indeed, it is the predictions of SBR supporters—that student and school accountability policies help students get better jobs and stay in college longer—that receive support. In addition, scores on tests that are not part of state accountability systems are higher in states with strong SBR policies. Thus, most students benefit from SBR policies. There are, however, some who lose out--those who would have graduated under the old rules but do not graduate because they cannot pass the tests. How large are these effects?

Effects on High School Graduation Rates: Our analysis of longitudinal data is presented in Figure 5. We found that the graduation rates of students with average or above average grades in 8th grade were not affected by state MCEs. However, students with C- grades in 8th grade were significantly (7.7 percentage points) less likely to get a high school diploma or a Graduate Equivalency Diplomas (GED) within 6 years when they lived in a MCE state. Graduation rates of students living in New York were no different from the graduation rates in states without MCEs. The share of students getting GEDs also went up in MCE and CBEEES states.

Figure 6 summarizes an analysis of state data on the ratio of diplomas awarded by public schools in 1998 to 8th grade public school enrollment in the fall of 1993. Figure 7 summarizes an analysis of state data on the ratio of diplomas awarded by public and private schools to the number of 17 year olds in the state in 1997 through 1999. States with higher non-academic course graduation requirements had significantly lower high school graduation rates. States with larger secondary schools had significantly lower graduation rates. None of the other policy variables had statistically significant effects. Nevertheless, point estimates for MCEs and CBEEES suggest that they probably lower graduation rates.

Let us now review the empirical findings regarding the efficacy of the different components of standards-based reform. States that reward schools for success and sanction schools that are failing had significantly higher achievement levels. These results are consistent with Grissmer et al's (2000) finding that the biggest gains in NAEP mathematics scores were in North Carolina and Texas—the two states that established the nation's most comprehensive systems of school and student accountability in the early 1990s. Students in MCE states were significantly (about 2 to 4 percentage points) more likely to attend college in 1993/94 and employers responded

to the their enhanced reputation by paying them 9 percent more. The effects of MCEs on achievement in 8th grade and test scores gains during high school were small and often not statistically significant. Curriculum-based external exit exam systems appear to have had by far the largest impacts on test scores. Achievement levels at the end of high school were roughly one grade level equivalent ahead of comparable states. Increases in the number of academic courses required for graduation also had substantial effects on learning during high school.

How can the Federal Government Help States Develop an Effective Standards-Based Reform Strategy for Secondary Schools

The federal government pays only a tiny portion of the costs of secondary education. How can it help reform secondary education and assist states in developing accountability mechanisms that produce better outcomes?

The first step has already been taken. The 2001 reauthorization of the Elementary and Secondary Education Act, the "No Child Left Behind" Act, requires states to test students at least once in grades 10-12 in reading, mathematics and science and to develop accountability systems based in part on that data. The implementation of this legislation will have profound effects on how standards-based reform is applied to high schools. The regulations for "No Child Left Behind", therefore, need to be informed by a vision of how standards based reform and high school reform should proceed. Consequently, this chapter will articulate a vision of how American high schools should be reformed based on the international and domestic evidence described in the first three sections of the paper. This vision is derived from and an extension of the administration's vision for the "No Child Left Behind" Act. As the discussion proceeds recommendations for those writing the regulations for "No Child Left Behind" will be presented in ***12 point bold Italics.*** New federal initiatives suggested by the argument will also be presented in ***12 point bold Italics.***

It is important to remember, however, that state governments are in charge here. They have constitutional responsibility for education and control the funding and the levers of authority that guide both K-12 and post-secondary education. It is their vision that will ultimately be implemented. Different states will make different choices. Some states use end-of-course exams to measure student achievement in high school [see Table 2]. Others use end-of-grade exams. Some have chosen to make high school graduation dependent on passing a state high school graduation test. Others have rejected high-stakes graduation tests. Michigan awards scholarships to students who

demonstrate proficiency on MEAP high school tests. Connecticut encourages employers and colleges to use state tests in their hiring and admissions decisions [see Table 3]. It would be a mistake for the federal government to attempt to use the regulations and grants for implementing "No Child Left Behind" (NCLB) to force all states to adopt a particular policy mix. The states are laboratories of democracy. Studying their contrasting experiences will teach us a great deal about what works and what doesn't.

The Optimal Design of Standards-Based Reform for High Schools:

Systems that hold high schools accountable for student learning are particularly difficult to design for five reasons. First, high schools have multiple goals. Some of these goals--achievement in core academic subjects and high graduation rates—apply to all schools and to all students. But others goals—speaking a foreign language, occupational competency, developing artistic talent and leadership skills—are goals that some students choose to pursue but many do not. If these specialist achievements are not recognized in the accountability system, administrators may be pressed to redirect resources away from these elements of the high school program. On the other hand, it is not easy to measure these student accomplishments comparably across schools. One would have to report both how many students were pursuing each goal and the standard achieved by these students. In applied technology, for example, one might report indicators such as (a) number of students taking two or more courses in each vocational specialization, (b) occupational skill certificates awarded to these students, (c) proportion of vocational students in school or employment six months after graduation, (d) proportion working or studying in the occupational field they studied in high school and (e) wage rate of those who are working full-time after high school. ***Implementation of the "No Child Left Behind" legislation should allow and indeed encourage states to include subjects other than English, mathematics and science in high school accountability systems.***

Secondly, measuring achievement in core academic subjects is more difficult for high school students than for elementary school students. Standards-based reform requires agreement at the state level on content standards for each subject, alignment of instruction with these content standards and alignment of assessments with both content standards and instruction. But unlike primary schools and middle schools, high schools lack a sequenced academic curriculum that everyone takes together. Students choose which math and science courses to take and when to take them. High achieving students often accelerate when they take math and science courses. How, then, does one design a challenging science test for tenth graders? Some take biology that year; others chemistry, physics, environmental science or earth science. Still others take no science.[19] A test covering all fields of science will inevitably be watered down and hold no one in particular accountable. It will be unlikely to improve peer norms in science classes. ***Separate assessments for each laboratory science course are a better way to bring***

accountability to high school science. Federal regulations should encourage (but not require) states to assess high school science courses individually rather than in one generic test. These exams would be administered at the end of each science course.

The third difficulty is that high school tests measure the cumulative result of ten to twelve years of schooling, not just what has been learned since the student entered high school. If students arrive in ninth grade not knowing how to read, it makes little sense to sanction the high school staff for a failure whose roots lie in the district's elementary and middle schools. This is one of the many reasons why school accountability systems need to measure value added and to give indicators of value added a central place in the definition of school quality. Since test scores from seventh and eighth grade will be available, indicators of value added can be constructed. The first step is to estimate models predicting high school test scores as a function of the student's 7th and 8th grade scores from a few years earlier. The prediction of this model for each student would be subtracted from the student's actual HST score and these deviations from the predicted score would be cumulated across all students in a school. If the mean deviation is positive, the high school is doing a better than average job. If the mean deviation is a large negative number, the school is failing to teach effectively. Unfortunately, many states currently lack the centralized student record keeping systems that are necessary to construct the value-added indicators described above. However, testing contractors have the information and expertise necessary to develop such indicators and this task should be added to the other tasks performed by the state's testing contractor. ***States will need time to decide how it's value added indicator should be defined, but NCLB regulations should require states to start the development process and to eventually incorporate such indicators in their accountability system.***

The fourth difficulty is that when a test is not part of a course's grade or important to the student in some other way, many high school students fail to put much effort into answering all the questions correctly and completely.[20] This doesn't pose a problem when a state's minimum competency high school graduation exam is used as the indicator of student achievement for high school accountability. But only 20 states currently have minimum competency exams. In most of the nation, tests that students have no reason to try hard on are the primary indicator of student achievement in school accountability systems. When this is the case, school ratings may reflect the school's success in getting students to try hard on state tests and rather than how much the students actually learned. This reduces the validity of high school tests as measures of true student achievement and tends to make their use in accountability systems problematic.

In the states that do not have high-stakes minimum competency exam graduation requirements, students can be induced to put effort into a school accountability test by giving them a stake in doing well. Where there are end-of-course exams or end-of-grade

exams in mathematics and English, the state exam can become one of the midterms or finals of the course. Another way to make the tests count is to persuade state universities and community colleges to use them in admissions decisions (in place of or supplementary to the ACT and SAT-1 tests) and for deciding whether entering students must take remedial courses. Still another approach is to award merit-based scholarships to students who demonstrate proficiency or high proficiency on them as Michigan and Ohio have done.[21]

The fifth problem in holding schools accountable is the low quality and low standards of many of the high school tests used in accountability systems. While student motivation is unlikely to be a problem when MCE scores are used in accountability systems, there are other problems. These tests determine who has not reached the minimum standard necessary to graduate. To avoid a political backlash, cut scores must be set low enough to insure that fewer than 10 percent of students are denied a diploma because they have been unable to pass one of the MCE tests. The performance level signaled by this cut score will be substantially below the standard we would like most students to achieve. To maximize the reliability of this high stakes classification and to shorten the test, test developers often omit difficult questions that marginal students are unlikely to answer correctly. As a result, scores obtained on most minimum competency exams do not describe the full range of student achievement the way Regents exams, AP exams, SAT-2s and teacher made exams do. Teaching to such an MCE would dumb down the curriculum for the majority of students who are not at risk of failing.

"No Child Left Behind" tries to prevent this problem from arising by adding a provision to the ESEA rules on state standards and assessment. The law requires that a state's academic standards include challenging student academic achievement standards that are aligned with the state's academic content standards; describe 2 levels of high achievement (proficient and advanced) that determine how well children are mastering the material in the state's academic content standards; and describe a third level of achievement (basic) to provide complete information about the progress of lower-achieving children toward mastering the proficient and advanced levels of achievement {Section 1111 (b)(1)(D)(ii)}.

Both the effects of standards-based reform and its long-term political viability depend on the quality and credibility of the exams used to measure student achievement. Consequently, implementation of the "No Child Left Behind" legislation should give priority to the development of high quality exams that are aligned with state learning standards in the subject and that require students to write essays, do multi-step math problems, conduct science experiments, etc. A great deal of work needs to be done. According to the Quality Counts 2002 report, six states have not yet developed content standards for high school mathematics and nine states have not developed content

standards for high school science. Criterion-referenced high school assessments aligned with state standards are not available in eight states for mathematics and in twenty-seven states for science. Only sixteen states use extended-response questions in their assessments of mathematics, science or social studies.

State departments of education (or their contractor) would develop the exams and the rubrics for grading extended answer portions of the exam and then train teams of teachers from the state to do the grading.[22] Each paper should be read at least twice. Grading exams collectively is invaluable professional development so as many teachers as possible should be recruited on a rotating basis. They should get a generous honorarium for the work. Grading should be done a week or so after testing so that students who fail the test can be put in an after-school program or retake the course in summer school. Quality exams take longer to develop, longer to take and longer to grade. Inevitably, they are more expensive.

How does the federal government discretely influence the choices the states make? The first step is to employ the bully pulpit. The President or the Secretary of Education should give a speech laying out his vision of how states should implement the testing provisions of the "No Child Left Behind" Act. At the beginning of the speech, he would say that states are the laboratories of democracy and he wants states to develop their own unique way of assessing student achievement. He would recommend a system with the following features:

- ***Tests that are comparable enough from year to year so that information is provided not only on how much Johnny knows, but how much he learned since last year. This is the kind of information that is needed to fairly assess a school's value added in the face of high rates of student turnover and large differences in the reading skills and family background of students entering a school.***
- ***The legislation requires that the tests provide "descriptive" and "diagnostic" information on the achievement of individual students. If diagnostic information is to be helpful, it needs to be reported back to the school soon after test administration so that remediation can begin immediately. It is unacceptable to wait until the end of the summer to get test results back.***
- ***Essays and extended response answers are an important part of the state's assessment and are graded by teachers, not by poorly trained temporary workers who have not completed college and are not residents of the state.***
- ***Test Security—Whenever stakes are attached to test results, test security has to be a concern. European high school exit exams, SATs***

ACTs and New York State Regents exams are all administered on the same day during a very small time window. New versions of the exam are constructed for each test administration. Similar security precautions are needed for state sponsored end-of-course exams and minimum competency exams.

We have to expect that many teachers will teach students how to handle the types of questions we put on the exam. The better the exam, the better the teaching will be. Consequently, NCLB language requiring states to develop "challenging student academic achievement standards" should be interpreted as meaning that the tests contain challenging content where students must do multi-step problems showing their work and explain their reasoning on science problems. All high school assessments should be peer-reviewed for alignment and quality. Implementation of the "No Child Left Behind" legislation should discourage states from buying cheap off-the-shelf tests that are poorly aligned with state learning standards in the subject. For example, all states include writing in their high school learning standards. NCLB regulations should require all states to develop an assessment of writing skills during high school that actually involves writing essays.

State university and community college systems need to work with state departments of education to improve the quality the state achievement exams for high school students and to develop ways of using these exams for admissions and placement purposes. ***The Department of Education should encourage such collaborations by establishing a grant program to fund them. The primary objective of the collaboration is to persuade the state's public institutions of higher education to use the end-of-course and high school graduation tests administered by the state's K-12 system when they make admissions and placement decisions. Community college and university systems that use their state's high school exit exams and end-of–course exams to help make admissions and placement decisions should have input into the design and revision of state tests. Since ninety percent of high school students aspire to go to college and seventy percent actually attend, it makes a great deal of sense to involve college teachers and administrators in the design of high school exams. These grants could help states develop ways to use high school graduation tests and end-of-course exams in deciding on admissions to state universities and colleges and determining placement of freshman in remedial or advanced courses.***

Optimal Design of Standards-Based Reform for High School Students:

Minimum Competency Exam (MCE) high school graduation requirements are the most common way that states make students accountable for learning. Studies of the

effect of MCEs have found that they increase college attendance and post high school earnings but have little effect on test score gains during high school and lower the probability that low GPA students get a high school diploma. A number of states appear to be following a strategy of driving their educational systems to higher standards by periodically revising their MCE in order to set progressively higher minimum standards.

Minimum Competency Exams create a High Stakes for a Few Students System: State tests determine or influence getting a diploma or promotion to the next grade but only a small minority of students are really at risk of being retained or being denied a diploma. One benefit of **High Stakes for a Few** is that it focuses school efforts on helping its most poorly prepared students. Critics of MCEs point to a number of problems with this approach:

a. There are other ways of getting schools to expend more energy on teaching lagging students. "Stakes for School systems" can be designed to accomplish this purpose.

b. Many perceive it to be **unfair** to, in Gary Orfield's words, **" punish"** students whose low test scores are the result [at least in part] of attending under funded poorly staffed schools. [I am not persuaded by Orfield's rhetoric because the benefits—higher wages and greater college attendance—of high school graduation tests are so large, they outweigh the losses experienced by the small number of students who fail to graduate because they cannot meet the standards. Nevertheless, initiatives that increase or modify the stakes for students need to be framed in a way that responds to this rhetoric.]

c. Most students put insufficient effort into their studies and avoid demanding courses, so incentives need to be strengthened for almost all students not just those who do poorly on tests.

d. Most students pass the MCE on the first try. Once they pass, the stimulus to studying and paying attention in class generated by the MCE goes away. Only in the minority of very troubled schools where the majority of students are at risk of failing the MCE is student culture likely to be changed by the high stakes test.

e. **Who is held accountable when students fail? Primarily the student. Possibly the principal.** In big high schools principals have limited ability to influence how their teachers teach. In most cases individual teachers are not considered responsible for how students in their class this term do on MCEs. Some MCEs are first administered in the fall. MCEs typically cover material studied in many different courses taught by different teachers. **When everyone is responsible for student performance, no one is responsible.**

f. The idea behind MCEs is that we fix the minimum graduation standard and then vary the time students devote to learning. By spending extra time at learning tasks, lagging students eventually achieve the higher standard. This is an attractive strategy. Fifteen of seventeen states with MCEs in 2001 required schools to provide remediation for students failing state MCE exams (Quality Counts 2002, p 77). Nevertheless, many school districts are not giving lagging students the extra learning opportunities after school and during the summer that they need to be successful.

g. MCE tests are designed to identify students whose achievement is so low they should not be awarded a diploma. To increase the reliability of this classification, test developers omit questions that the marginal students are unlikely to be able to answer. If regular instruction comes to focus on preparing students for the MCE test, the majority of the students who are not at risk of failing will be getting a diluted and undemanding curriculum.

MCE graduation requirements tend to be politically controversial. Raising the bar often seems impossible because failure rates on pilot administrations of new MCEs are typically very high. State education leaders in Arizona, Wisconsin and Massachusetts have recently been forced to either postpone the MCE graduation requirement or reduce the stringency of the testing requirement. Whatever ones personal view of how the benefits of MCEs compare to their costs, it is clear that the political culture of many states rules out this policy option. If a state does not want to make the high school diploma contingent on passing a MCE test, what can it do to induce high school students to take learning seriously? The next subsection describes a series of powerful ways of giving students a bigger stake in learning without imposing high stakes negative consequences on them if they are unsuccessful

Moderate Stakes for Everyone should be the objective, not high stakes for the few. A number of ideas for generating moderate rewards for learning are described below. While states with no MCE have the greatest need to implement these approaches, these proposals can improve motivation and student culture in MCE states as well.

1. **Make the consequences of doing poorly on state tests less draconian.** Retention should be reserved for only the most egregious cases and only after extra time remediation efforts have been tried and failed. Instead of being retained, students who are falling behind should be **required** to participate in:

    * **After-School Programs**
    * **Saturday School Programs**
    * **Summer School Programs**

Consequences such as these are likely to be at least as strong an incentive to study hard as the threat of retention. Yet they do not "punish" the student, they help remedy the poor reading skills etc. that are the source of the problem. Requiring students to participate in extra-time learning opportunities should not depend solely on scores on state tests. Teachers should also have input in a decision made either by the principal or a committee.

The Administration should propose a further major expansion of the program of grants to school districts to provide expanded after-school and summer school opportunities for children who are not doing well in school. The Education Secretary and the President should encourage school districts that are "ending social promotion" to give lagging students at least one full year of after-school and summer school remediation before holding a student back. States should be encouraged to pass laws giving school districts the authority to require students who are falling behind to attend school during the summer.

2. ***The administration should push for a big expansion in the number of students taking Advanced Placement (AP) and International Baccalaureate (IB) courses and examinations.[23] This can be accomplished by funding summer institutes for the teachers of AP and IB courses and by negotiating a reduction in the fee for taking the AP and IB examinations. The U.S. Department of Education should study and evaluate state efforts to offer internet-based AP courses to students attending small high schools and fund enhancements and quality improvements of these courses. Grants should be given to states that have developed exemplary courses so that students from other states can take the course for a nominal fee. Private non-profit organizations that have developed exemplary Internet courses should also be allowed to compete for these grants.***

3. **Graduated Rewards for Doing Well on State Tests**. The rewards should not be large amounts of money for exceeding a cutoff. They should be graduated and based on absolute performance, not performance relative to the other students in the school. All of these ideas have already been implemented by a few states [see Table 3]. Additional states should implement with these policies.

   - **Scores on state tests should be part of the final grade in the course.** This will require that state tests be quickly graded before the end of the school year.

   - **Scores on state tests should be on the high school transcript**

   - **Differentiated diplomas or honors certifications on the existing diploma.** Student eligibility for honors diploma certifications should depend (at least in

part) on their performance on external exams and possibly the rigor of the courses being taken. They should not depend on an unweighted GPA. If a MCE is in place, students who fail the MCE but get the requisite number of Carnegie units should get a certificate of completion and be allowed to walk across the stage.

- **Merit Scholarships** similar to the Michigan Merit Award that are based on students' grades on a battery of the state's external exams. They should be awarded at assemblies attended by parents. These merit scholarships would not have to be for large amounts of money. Better to award lots of them than award large stipends. The size of the award could depend on financial need. This would compensate for the advantages that students with wealthy parents have in the competition for these scholarships. ***Once a state has implemented a set of reliable high quality assessments aligned with state content standards for grades 9 through 12, the federal government should offer to match state funds allocated to a state merit scholarship program that selects awardees largely on the basis of scores on the state assessments. Students in private high schools should be eligible for these awards if the bulk of students at the school participate in the state's testing program. In the first year of the state's merit scholarship program the federal contribution might be formula based [e.g. $500 per high school graduate]. States would structure the eligibility rules so that roughly one-third of high school graduates would be able to receive the merit scholarship in the first year. The amount of the award would vary with achievement level and financial need, but everyone would get a minimum of $500. Thus, the maximum award for low-income students with very high scores might be as high as $10,000. Over time achievement will improve and the share of graduates meeting the standard and receiving the scholarship will rise as well. The federal contribution would increase proportionately.***

- **Recruit and publicize employers who promise to pay students with the honors certifications a higher wage.** Connecticut has done exactly this.

- **Persuade State Colleges and Universities to announce that they use grades on state tests in admission and placement decisions.**

4. America's premier high stakes tests, the SAT-I and ACT, are not comprehensive measures of learning during high school.[24] The energy that students devote to cracking the SAT-1 would be better spent reading widely and learning to write coherently, to think scientifically, to analyze and appreciate great literature and to converse in a foreign language. These are the true objectives of a high school education. The high stakes attached to the SAT-1 and the ACT, however, tend to

direct student energy away from developing these important skills and weakens the ability of teachers to set high standards themselves.

Colleges should redirect the energy of high school students towards our true educational objectives by dropping the SAT-1 and ACT tests and replacing them with state sponsored curriculum-based end-of-course exams like New York State's Regents exams and/or national subject specific achievement exams like the SAT-2, Advanced Placement and International Baccalaureate exams (Kirst 2001). Changing admissions criteria in this way will help convince students, parents and school administrators that better teaching, more challenging courses and higher achievement will be perceived and rewarded by the colleges and universities.

The Secretary of Education should give a speech supporting the proposal by the President of the University of California, Richard Atkinson, to substitute achievement exams like the SAT-2, AP exams and state end-of-course exams for the SAT-1 and ACT exams in admissions and class placement decisions of California's state colleges and universities. In order to accelerate the transition from the SAT-1 to state developed achievement tests, the Office of Education Research and Improvement should fund studies that (a) compare the validity of state achievement tests, SAT-2, SAT-1 and ACT tests in predicting college grades and degree completion and (b) empirically compare the scoring standards of achievement exams from neighboring states.

The Department of Education should also make grants to collaborations between state community college systems, state university systems and state education departments to develop ways to use state high school graduation tests reflecting high standards (e.g. MCAS, MEAP, the SOLs, etc.) and end-of-course exams in deciding on admissions to state universities and colleges and for placement of freshman in remedial or advanced courses in community colleges, technical institutes and state universities. Funding priority should go to states that establish a permanent institutional mechanism for regular discussions between K-12 and higher education regarding the coordination of high school graduation requirements and tests with college admissions and placement tests and requirements.

High schools should hold **all** students to higher standards. Poorly prepared students need to be told of their deficiencies early in high school when there is time to remedy them. If that is done, the share of college freshman with the skills and knowledge necessary to succeed will rise and many more will realize their goal of getting a bachelors degree.

5. End-of-Course Exams (EOCEs) should be the core of accountability for high

school students. The regression analysis of state NAEP test scores and dropout rates summarized in section 3 of this paper found that end-of-course exams had more positive effects on learning and retention than high stakes MCEs and the no/low stakes end-of-grade exams. Why? Because:

a. **Responsibility for student performance on a particular exam is focused on just one or a small group of teachers.**

b. **The classroom culture is improved because everyone is taking the same exam and it will be part of the student's grade in the course.** EOCEs signal the full range of achievement in the subject; so everyone has an incentive to study harder in order to do better on the test; not just the students at risk of failing the course.

c. **Student attitudes towards that teacher are improved** because she becomes a coach who helps the class succeed on the state exam. Her role shifts from being a judge towards being a mentor. New York State has an EOCE system. Connecticut, Massachusetts and New Jersey do not. Contrasting NY and its neighbors allows us to test this assertion. Surveys of 35,000 students in these states by the Educational Excellence alliance found that attitudes toward teachers were more positive in New York. When students were asked what motivated them to study hard, New Yorkers were 30 percent more likely to respond "to please or impress my teacher," 17 percent more likely to say 'my teachers encourage me to work hard.' and 14 percent more likely to say "the teacher demands it." New York students were also significantly more likely to say "my teachers grade me fairly", "my teachers maintain good discipline in the classroom" and that classes are "interesting."

d. **Student peer support for studying and classroom engagement increases. Peer support of disruptive students decreases.** New York students were 10 percent more likely to say, "My friends think it is important for me to do well in [science, math, English] at school." They were nearly 25 percent more likely to be annoyed when "other students talk or joke around in class" or "try to get the teacher off track." In addition New York students were significantly more likely to say they were motivated by a desire to learn the material and more likely to report they were interested in what they were studying and more likely to talk with their friends outside of class about what they were studying. The better attitudes translated into better behavior. New York students spent significantly more time studying for history exams, more time doing homework and did a larger share of the homework that was assigned. They also paid closer attention in class and contributed to class discussion more frequently.

e. **EOCEs assess more difficult material.** Since EOCEs are supposed to measure and signal the full range of achievement in the subject, they contain more difficult questions and problems. This induces teachers to spend more time on cognitively demanding skills and topics

f. **Students take the course when they are ready for it.** Alignment between instruction and the exam is maximized.

g. **Teachers grade the exam.** Grading exams with essays and other constructed response questions is a very effective form of professional development. In NY, teachers participate in the grading of their own student's exams, so they get good feedback on where their teaching failed.

Table 1: Characteristics of Secondary Education Systems

	Enrolment Rate		Upper Sec. School	Learning Index 4th to 8th Grade				End of Secondary		Lower Sec. Teacher	Up. SS Salary	Teacher Hours	Salary Per	Expend Perstud	Teacher Comp	Student Hours	%Absent
	Age 16	Age 17	Grad. Rate	Math	Science	Reading		Math	Science	Salary In $	/GDP percap	per Year	Teach Hours	GDP Percap	% of total	Per Year	8th Grade
Australia	97	81	—	121	127	—	499	522	527	$36,175	1.6	802	45	25	63	1027	7.1
Belgium	94	93	84	—	—	-14	539	—	—	$27,932	1.55	685	40	29	78	1057	4.1
Canada	99	83	72	133	130	+26	498	519	532	—	—	—	—	—	64	—	5.4
Denmark	93	82	—	—	—	+50	514	547	509	$31,000	1.6	572	48	28	53	930	3.6
Finland	89	93	89	—	—	-10	*520*	—	—	$27,942	1.3	457	58	25	60	855	—
France	95	88	87	—	—	+7	498	523	487	$29,615	1.3	620	47	31	—	975	3.7
Germany	96	91	93	—	—	+21	476	495	497	$38,640	1.9	710	53	28	—	921	4.1
Hungary	97	85	90	127	175	+44	504	483	471	$11,066	1.0	555	20	21	—	902	4.4
Italy	78	73	—	—	—	+3	*479*	476	475	$25,773	1.2	612	42	29	73	1105	—
Japan	96	94	96	148	140	—	572	—	—	$41,201	1.7	—	—	24	—	875	2.1
Korea	96	90	90	137	105	—	591	—	—	$39,921	2.7	494	80	24	—	867	0.9
Netherlands	96	85	93	103	150	+29	519	560	558	$31,380	1.9	910	34	23	—	1067	3.0
Norway	94	93	—	138	150	-9	483	528	544	$23,879	…	558	39	26	—	855	3.4
Portugal	84	81	56	115	165	+35	416	—	—	$26,288	1.7	571	42	29	—	878	5.0
Spain	85	73	67	—	—	+2	452	—	—	$32,144	2.0	545	59	27	75	957	3.2
Sweden	98	97	79	—	—	+7	497	559	552	$23,896	1.2	—	—	27	44	741	4.4
Switzerland	90	85	84	—	—	+21	519	540	523	$51,361	2.1	768	60	42	72	—	—
U. K.	81	66	—	130	149	—	482	—	—	$38,010	1.7	798	48	23	51	720	6.0
United States	**84**	**74**	**74**	**93**	**113**	**-15**	**472**	**461**	**480**	**$32,713**	**1.1**	**954**	**34**	**25**	**57**	**980**	**5.4**

Sources: OECD, Education at a Glance , 2000, pp. 95, 103, 136, 147, 215, 237. Warwick Elley, How in the World do Students Read? 1972, p 108-9; Beaton, Albert et al. (1996) Mathematics [Science] Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study. CSTEEP, Boston College, Boston MA. //:http/wwwcsteep.bc.edu/TIMSS. Scores in italics are for 8th graders in TIMSS-R.

33

32

Table 2: End-of-Course Examination Systems

State	Yr **Announced**	**Subjects** (year first administered)	Score On Tran-script	Part of Cours e Grade	Teach-ers Grade Exam	Honors Diplom a based on EOCE	Whe n MCE bega n	EOCE can substitute for MCE	Other Rewards for Student Achievement
New York	1865	English, Math (3), Biology, Chemistry, Physics, U.S. History, World History, Latin, Foreign Languages, Intro. to Occupations	Yes	Yes	Yes	Yes (40%)	1979	about 1992	In 1950s Scholarships were based on Regents exams. Use in teacher assessment is a local option. Becomes primary HS Grad. Test after 2000-03
North Carolina	1984	Algebra 1 & Biology (1987), Algebra 2 & US History (1988), Chem. & Geometry (1989) Eng. 1, Physics & Soc. Studies (1990-1),	Yes	Most (after 2000 25%)	Yes	2003	1980	No	State Tests at earlier grades influence retention decisions
California	1983	Algebra I & Geometry (1987), U.S. History & Economics (1990), Biology & Chemistry (1991), Coord. Science (1994), Writing (1996), Civics (1997), Liter. & H.S. Math (1998), Physics & Spanish (1999)	Yes	No	No	Yes (1%)	2004	No	State tests at earlier grades influence retention decisions
Texas	1992	Biology (1995), Algebra I (1996), US History & English (1999)	Yes	Most (Req. in Future)	?	No	1987	2000	Scholarships based on course rigor and family income. State tests at earlier grades influence retention decisions.

Tennessee	1992	Alg. 1, Biology & English II (2001) Algebra 2, Geometry, Eng I (2002), US Hist, Chem. & Physics (2003)	Yes	Yes	?	No	1985	2005	Becomes the HS graduation test in 2005. Current Honors Diploma based on GPA.
Maryland	1995	English 1, Civics, Algebra, Geometry & Biology (2001),	Yes	?	?	No	1983	2007	Becomes the HS graduation test in 2007. Honors Diploma based on rigorous courses and GPA starts in 1998
Mississippi	1994	Algebra & US History (1997), Biology (1998)	?	?	?	No	1989	No	Merit Scholarship based on GPA and ACT scores.
Virginia	1996	English, Algebra 1 & 2, Geometry, Earth Science, Biology., Chemistry, US Hist, World Hist. (1998)	Yes	some	?	Yes	1981	2004	Becomes HS graduation test in 2004. State tests at earlier grades influence retention decisions
Oklahoma	1999	English & US History (2000), Math & Biology (2001)	Yes	No	No	No	none	No	State Univ and Employers encouraged to use EOCE
Arkansas	1997	Math (1999), English. (2002), Science & History (2004)	Yes	No	No	No	none	No	State tests at earlier grades influence retention decisions

Table 3: Statewide Examinations that Determine Eligibility for Honors Diplomas or Scholarships

State	Yr Announced	Subjects (year first administered)	Score On Tran-script	Part of Grade	Do Teachers Grade Exam?	Honors Diploma based on EOGE	When HSGT Req. begins	Month of First Admini–stration	Other Rewards for Student Achievement
Ohio	1987	12th Grade tests in Reading, Science, Math and Civics (1994-6)	Yes	No	No	Yes In part	1994	Feb.	$500 scholarship based on EOCE. Honors Diploma req. rigorous courses & either GPA gt 3.5 or 12th grade exams or ACT.
Conn.	1991	10th Grade tests in English, Math and Science (1994)	Yes	No	No	Yes	none	May	
Michigan		11th Grade tests in Math, Reading, Science and Writing (1997), Social Studies (1999)	Yes	No	No	1996 (Subject–by – subject)	none	March	Beginning with 1999 graduates a $2500 scholarship is awarded based on EOG exams.
Penn.	1991	11th Grade tests in reading, writing, math (1999)	some	No	No	2003	none		
Oregon	1991	10th grade tests in English & Math (1996), Science (1999), Social Studies (2003)	Most. Expect an Increas e	Some	teachers blind	2001 (propose Subj–by – subject)	none	Feb(PA) April (MC)	Certif. of Initial Mastery based on English & math in 2001, add Science in 2002, add Arts in 2003, add 2nd Lang in 2005, & Soc. Studies in 2006.
Illinois	1997	11th Grade tests in reading, writing, math, science & social science (2001)	Yes (Req-uired)	No	No	2002 (Subj–by –subject)	none		
Phase in of High School Graduation Test									
Mass.	1993	10th Grade tests in English, Math, & Science (1998)	No Temp	No	No	2000	2003	May	On 3/28/00 State Bd of Ed decided to move up the first class getting Certificate of Mastery to 2000. Based on either EOG scores, AP or SAT 2's.
Wisconsin	1997	10th Grade tests in reading, writing, math, science & social science (2002)	Yes	No	No	No	2004	Spring	1997 legislation with HGST repealed in 1999. Local Districts will set graduation standards based on HGST and other indicators of student achievement
Indiana	1993	10th grade tests in English and mathematics (1997)	Most	No	No	No	2000	Sept.	May also meet graduation requirement by getting a C or better in all Core 40 college prep courses or demonstrate 9th grade achievement in other way. Honors Diploma based on Curriculum

References

Bishop John H., "Are National Exit Examinations Important For Educational Efficiency?" Swedish Economic Policy Review, Vol. 6, #2, Fall 1999, 349-401.

Bishop, John, Ferran Mane, Michael Bishop and Joan Moriarty. (2001) "The Role of End-of-Course Exams and Minimum Competency Exams in Standards-Based Reforms." Brookings Papers on Education Policy, edited by Diane Ravitch, (Washington, DC: The Brookings Institution), 267-345.Michael

Ehrenberg, Ronald and Brewer, Dominic. "Did Teacher's Race and Verbal Ability matter in the 1960's? Coleman Revisited." Ithaca, NY: Cornell University, School of Industrial and Labor Relations, 1993, 1-57.

Elley, Warwick, How in the World do Students Read?, The Hague, The Netherlands: International Association for the Evaluation of Educational Achievement, 1992.

Ferguson, Ronald. Racial Patterns in How School and Teacher Quality Affect Achievement and Earnings. Cambridge Mass: Kennedy School of Government, Harvard University, 1990.

Fortner, Cmaeron. "Who's scoring those high-stakes tests? Poorly trained temps." The Christian Science Monitor, September 18, 2001, www.csmonitor.com/2001/0918/p19s1-lekt.htm.

Frederick, W. C. "The Use of Classroom Time in High Schools Above or Below the Median Reading Score." Urban Education 11, no. 4 (January 1977): 459-464.

Frederick, W.; Walberg, H.; and Rasher, S. "Time, Teacher Comments, and Achievement in Urban High Schools." Journal of Educational Research 73, no. 2 (Nov.-Dec. 1979): 63-65.

Gamoran, A and Barends, M. (1987) "The Effects of Stratification in Secondary Schools: Synthesis of Survey and Ethnographic Research." Review of Education Research. Vol. 57, 415-435.

Goodlad, J. A Place Called School. New York: McGraw-Hill, 1983.

Grissmer, David and Ann Flanagan, Jennifer Kawata and Stephanie Williamson. Improving Student Achievement: What NAEP test scores tell us. Rand Corporation, 2000, 1-271.

Hanushek, E. A. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro-data." American Economic Review, 61(2), 1971, 280-288.

Hayward, Ed. "Dramatic Improvement in MCAS scores" Boston Herald, Oct. 16, 2001. www.bostonherald.com/news/local_regional/mcas10162001.htm

International Assessment of Educational Progress, "Proficiency Scores and Graphs for All Populations." Report # 11, June 1992, 1-13.

Kirst, Michael. "State Education Standards and Admission/Placement Requirements." Paper presented at the College Board Conference, "New Tools for Admission to Higher Education," December 2001, The Bridge Project, Stanford University, 1-20.

Klein, M F.; Tyle, K. A.; and Wright, J. E. "A Study of Schooling Curriculum." Phi Delta Kappan 61, no. 4 (December 1979):244-248.

Longitudinal Survey of American Youth. "Data File User's Manual" Dekalb, Ill: Public Opinion Laboratory, 1988.

Monk, David. "Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement." Department of Education, Cornell University, 1992, 1-51.

National Center for Educational Statistics. The Digest of Education Statistics: 2000. Wash. D.C.: US Department of Education, 2000.

National Center for Educational Statistics. The Condition of Education: 2000. Vol. 1, Wash. D.C.: US Department of Education, 2000.

National Center for Educational Statistics. Occupational and Educational Outcomes of Recent College Graduates 1 year after Graduation: 1991. NCES 93-162, Wash. D.C.: US Department of Education, 1993.

Olson, Lynn, "States turn to curriculum-based tests." Education Week, June 5, 2001

Powell, Arthur; Farrar, Eleanor and Cohen, David. The Shopping Mall High School. New York, New York: Houghton Mifflin, 1985.

Sizer, Theodore R. Horace's Compromise: The Dilemma of the American High School. Boston: Houghton Mifflin, 1984.

Strauss, R.P. and Sawyer, E.A. "Some New Evidence on Teacher and Student Competencies." Economics of Education Review, 5(1), 1986, 41-48.

Tucker, Marc. "The Roots of Backlash: A Midterm Assessment of the Standards and Accountability Movement." Education Week , Vol. 21, #16, Jan. 9, 2002, pp 76 & 42.

Wiley, David E. "Another Hour, Another Day: Quantity of Schooling, a Potent Path for Policy." In Schooling Achievement in American Society, edited by William H. Sewell, Robert M. Hauser, and David L. Featherman. New York: Academic Press, 1976.

Endnotes

---

[1] Organization of Economic Cooperation and Development, Education at a Glance , 2000, p. 136, 147.

[2] For reading we make comparisons by subtracting a country's mean score for 9 year olds from its mean score for 13 year olds. A negative number indicates that a country's students have learned less in the interim than other countries in the study. A positive number indicates they learned more. IEA reading tests were each given arbitrary international means of 500 and standard deviations of 100 .

[3] Warwick Elley, How in the World do Students Read? (The Hague, International Association for the Evaluation of Educational Achievement, 1972) p. 108-9;

[4] Richard Ingersoll, *Out of Field Teaching and Educational Equity* NCES 96040, (Washington, DC: National Center for Educational Statistics, 1996).

5. OECD, Education at a Glance, Paris, 2000. P. 119 & 103.

[6] Laurence Steinberg, Bradford Brown, and Sanford Dornbusch, *Beyond the Classroom* (New York: Simon and Schuster, 1996), pp. 145-146.

[7] James A. Kulik and Chen-Lin Kulik, "Effects of Accelerated Instruction on Students," *Review of Educational Research*, Vol. 54 No. 3 (Fall 1984), pp. 409-425; David Monk, "Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement," *Economics of Education Review*, Vol. 13 No. 2 (1994), pp. 125-145. and John H. Bishop, "Incentives to study and the organization of secondary instruction." *Assessing Educational Practices*, eds. William Baumol and Becker (Cambridge, Mass.: MIT Press, 1996), pp. 99-160.

[8] Longitudinal Survey of American Youth, "Data File User's Manual" Q. AA37N.

[9] Interview with counselor at a wealthy suburban school, August 1997

[10] Ward, "A Day in the Life," *N.Y. Teacher* (Albany, New York, January 1994).

[11] Peter D. Hart Research Associates, "Valuable Views: A public opinion research report on the views of AFT teachers on professional issues" (Washington D.C.: American Federation of Teachers, 1995), pp. 1-24.

[12] John H. Bishop, (1996) "The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning." International Journal of Education Research; John H. Bishop, "The Effect of National Standards and Curriculum-Based External Exams on Student Achievement." American Economic Review, May 1997, Similar results were obtained by Ludger Wößmann, "Schooling Resources, Educational Institutions, and Student Performance: The International Evidence," Kiel Working Paper No. 983, (May 2000) Kiel Institute of World Economics, Germany, <http://www.uni-kiel.de/ifw/pub/kap/2000/kap983.htm> 1-88. I have redone the analysis of TIMSS data incorporating TIMSS-R data on student performance in 1999. This revised analysis finds the effects of CBEEES to be larger than in the earlier studies published by IJER, AER and the Swedish Economic Policy Review.

[13] Minimum competency exams are additions to, not a replacement for standards set by teachers. In a MCE regime, teachers continue to control the standards and assign grades in their own courses. Students must still get passing grades from their teachers to graduate. The MCE regime imposes an additional graduation requirement and thus cannot lower standards (Costrell 1998). The Graduate Equivalency Diploma (GED), by contrast, offers students the opportunity to shop around for an easier (for them) way to a high school graduation certificate. As a result, the GED option lowers overall standards. This is reflected in the lower wages that GED recipients command. Stephen V. Cameron and James J. Heckman, "The Nonequivalence of High School Equivalents" Working Paper # 3804 (Boston, Mass.: National Bureau of Economic Research, 1991).

[14] "Quality Counts," *Education Week*, January 11, 1999, p.87.

[15] "Quality Counts," *Education Week*, January 11, 1999, p.93.

[16] In 1996 only 4 of the 17 states with MCEs targeted their graduation exams at a 10th grade proficiency level or higher. Failure rates for students taking the test for the first time varied a great deal: from a high of 46% in Texas, 34 % in Virginia, 30% in Tennessee and 27% in New Jersey to a low of 7% for Mississippi. However, since students can take the tests multiple times, eventual pass rates for the Class of 1995 were much higher: 98% in Louisiana, Maryland, New York, North Carolina and Ohio; 96 % in Nevada and New Jersey, 91% in Texas and 83% in Georgia. American Federation of Teachers, *Making Standards Matter:1996* (Washington, DC: American Federation of Teachers, 1996) p. 30.

[17] One can also see in figure 2 that in most of the United States students with A averages do not get better jobs immediately after high school than C students. In fact when one holds college attendance constant, they tend to earn considerably less. Because Regents exam scores are part of student grades and appear on high school transcripts (thus signaling who is taking a more rigorous curriculum), we checked to see whether rewards for academic achievement were greater in New York State than elsewhere in the nation. This hypothesis was confirmed.

[18] The cross section analysis of state data on NAEP test scores and dropout rates included controls for the percent of children living in poverty, parental education, percent foreign born, the percent of public school students who are African-American and the percent who are Hispanic.

[19] In 2000 only seventeen states required students to take at least three science courses to graduate from high school. Digest of Education Statistics 2000, Table 154.

[20] This observation is based on interviews with the directors of the testing and accountability divisions in Manitoba and New Brunswick Canada and the large increases in student performance that occurred in New Brunswick, Massachusetts, Michigan and other states when no-stakes tests become moderate or high-stakes tests (Ed Hayward, "Dramatic Improvement in MCAS scores" Boston Herald, Oct. 16, 2001). Experimental studies confirm the observation. In Candace Brooks-Cooper master's thesis, a test containing complex and cognitively demanding items from the NAEP history and literature tests and the adult literacy test was given to high school students recruited to stay after school by the promise of a $10.00 payment for taking a test. Students were randomly assigned to rooms and one group was promised a payment of $1.00 for every correct answer greater than 65 percent correct. This group did significantly

better than the students in the other test taking conditions, one of which was the standard try your best condition. Candace Brooks-Cooper, 1998.

[21] Michigan gives a one-year $5000 scholarship to all students who score at the proficient or above level on MEAP high school tests in reading, mathematics, science and writing. Since instituting the scholarship program in 1999, test boycotts have ended and the number of low scoring students has fallen a great deal. The proportion of students achieving proficiency has risen substantially and the number of seniors planning to go to college has risen as well (Bishop, 2001).

[22] Non-teachers (generally college students) who do not live in the state grade the extended answer portions of most state tests. A Stanford graduate who worked for one of the testing companies grading the essays completed by 8th graders all over the nation described his colleagues as "temporary employees who had little respect for—and minimal investment--in their jobs." Cameron Fortner, "Who's scoring those high-stakes tests? Poorly trained temps." The Christian Science Monitor, September 18, 2001, www.csmonitor.com/2001/0918/p19s1-lekt.htm.test

[23] The number of students taking Advanced Placement (AP) examinations has been growing at a compound annual rate of 9 percent per year. In 1999 686,000 students, about 11 percent of the nation's juniors and seniors, took at least one AP exam. Despite this success, however, 44 percent of the high schools do not offer even one AP course and many others allow only a tiny minority of their students to take these courses. College Board, "More Schools, teachers and students accept the AP challenge in 1998-99," (New York, Aug. 31, 1999), pp. 1-8.

[24] The SAT-I and the ACT fail to assess most of the material--economics, civics, literature, foreign languages and the ability to write an essay--that high school students are expected to learn. The SAT-I leaves history and science out as well. The ACT's science and history subtests are very short and not linked to specific curricula. They are as much a reading test as a test of content knowledge in science and history.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").